Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data

Sairanen, V.

2018-11-01

Sairanen , V , Leemans , A & Tax , C M W 2018 , ' Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data ' , NeuroImage , vol. 181 , pp. 331-346 . https://doi.org/10.1016/j.neuroimage.2018.07.003

http://hdl.handle.net/10138/310910
https://doi.org/10.1016/j.neuroimage.2018.07.003

NeuroImage 181 (2018) 331–346

ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data

V. Sairanen[a,b,*], A. Leemans[c], C.M.W. Tax[d]

[a] Department of Physics, University of Helsinki, Helsinki, Finland
[b] HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[c] Image Sciences Institute, University Medical Center Utrecht, Utrecht, Netherlands
[d] Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, United Kingdom

A B S T R A C T

The accurate characterization of the diffusion process in tissue using diffusion MRI is greatly challenged by the presence of artefacts. Subject motion causes not only spatial misalignments between diffusion weighted images, but often also slicewise signal intensity errors. Voxelwise robust model estimation is commonly used to exclude intensity errors as outliers. Slicewise outliers, however, become distributed over multiple adjacent slices after image registration and transformation. This challenges outlier detection with voxelwise procedures due to partial volume effects. Detecting the outlier slices before any transformations are applied to diffusion weighted images is therefore required. In this work, we present i) an automated tool coined SOLID for slicewise outlier detection prior to geometrical image transformation, and ii) a framework to naturally interpret data uncertainty information from SOLID and include it as such in model estimators. SOLID uses a straightforward intensity metric, is independent of the choice of the diffusion MRI model, and can handle datasets with a few or irregularly distributed gradient directions. The SOLID-informed estimation framework prevents the need to completely reject diffusion weighted images or individual voxel measurements by downweighting measurements with their degree of uncertainty, thereby supporting convergence and well-conditioning of iterative estimation algorithms. In comprehensive simulation experiments, SOLID detects outliers with a high sensitivity and specificity, and can achieve higher or at least similar sensitivity and specificity compared to other tools that are based on more complex and time-consuming procedures for the scenarios investigated. SOLID was further validated on data from 54 neonatal subjects which were visually inspected for outlier slices with the interactive tool developed as part of this study, showing its potential to quickly highlight problematic volumes and slices in large population studies. The informed model estimation framework was evaluated both in simulations and in vivo human data.

1. Introduction

Diffusion MRI (dMRI) is sensitized to the microscopic motion of particles (Basser et al., 1994; Stejskal and Tanner, 1965) and is extensively being used to study brain connectivity and tissue microstructure in case of normal development, disorders, and training (Jones, 2008; Langen et al., 2012; Odish et al., 2015; Baum and Stevenson, 2016; Hämäläinen et al., 2017). Diffusion weighted images (DWI) are, however, affected by a range of artefacts, including image misalignment due to subject motion and geometrical distortions due to field inhomogeneities and eddy currents (Tournier et al., 2011; Tax et al., 2016; Andersson and Skare, 2010; Pierpaoli, 2010; Heemskerk et al., 2013; Kennis et al., 2016). Commonly used correction strategies for these artefacts are based on the post-acquisition registration and geometrical transformation of the images, and are integrated into various software tools for dMRI processing (Pierpaoli and Walker, 2010; Jezzard et al., 1998; Mangin et al., 2002; Nielsen et al., 2004; Rohde et al., 2004; Leemans et al., 2009).

In addition to geometrical artefacts, intensity errors that are nonrelated to the microscopic motion of particles can impede an unequivocal characterization of the true physical diffusion process (e.g. Vos et al., 2017). For example, subject head motion during the acquisition of a DWI slice can induce artificial signal decreases whereas hardware issues could also lead to erroneous signal increases (Jones and Cercignani, 2010; Le Bihan et al., 2006) which cannot be ameliorated with geometrical transformations, even though the error sources for both artefacts could be the same. This is especially harmful for clinical datasets with relatively few directions per shell as artefactual measurements can have high effect on the final results. Although multiple tools have been developed to detect intensity errors, the importance and complexity of this problem has recently been re-emphasized (Andersson et al., 2016). Voxelwise detection strategies (Cook et al., 2006; Lucas et al., 2010; Pierpaoli and

* Corresponding author. HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, POB 340, Haartmaninkatu 4, 00029, Helsinki, Finland.

E-mail address: viljami.sairanen@helsinki.fi (V. Sairanen).

https://doi.org/10.1016/j.neuroimage.2018.07.003
Received 23 October 2017; Received in revised form 22 May 2018; Accepted 2 July 2018
Available online 5 July 2018

Walker, 2010; Pannek et al., 2012; Chang et al. 2005, 2012; Collier et al., 2015; Tax et al., 2015; Mangin et al., 2002) based on robust estimators are typically performed after correction for geometric distortions, but the geometric transformations cause slicewise intensity artefacts to spread over multiple locations. Slicewise intensity errors, in particular, become distributed over multiple adjacent slices and appear as gradual intensity bands within the resampled DWIs. The edge regions of such gradual intensity bands are particularly problematic for these voxelwise robust estimators due to partial volume effects; the local intensity is an interpolation between normal and outlier data due to the transformation and resampling, and as a result such partial volume outliers could fall within the normal noise variation. Slicewise intensity outliers must therefore ideally be detected prior to any DWI geometric transformation (Morris et al., 2011; Andersson and Sotiropoulos, 2016; Marami et al., 2016; Marami et al. 2016, 2016), which is not generally adopted (e.g. Lauzon et al., 2013). In addition, using voxelwise strategies to detect slicewise outliers depends on the choice of voxelwise diffusion signal model or representation, and does not fully exploit the information present in the signals of the rest of the slice.

Several strategies have been proposed specifically for slicewise outlier detection (Zhou et al., 2011; Lauzon et al., 2013; Liu et al., 2010; Oguz et al., 2014; Jiang et al., 2006; Li et al., 2013; Andersson et al. 2016, 2017; Andersson and Sotiropoulos, 2016; Pannek et al., 2015; Scelfo et al., 2012), which are either based on the slicewise comparison of intensity metrics within the same DWI or on using slicewise information across DWIs. The algorithm of Liu et al. (2010); Oguz et al., 2014 (as implemented in DTIPrep) does within-DWI detection based on the normalized cross-correlation of adjacent slices and rejects the whole DWI volume if it contains multiple outliers. However, an affected DWI may still contain useful information, and ideally only the corrupted slices should be detected and their effect on subsequent analysis be reduced. Jiang et al. (2006), Li et al. (2013) and Marami et al. (2016a,b) based their detection to a morphological closing operation (implemented in DTI Studio) to calculate voxelwise indices that are used to guide the model estimation. Both of the aforementioned algorithms can result in false positives in the case of multiple adjacent or interleaved outliers. Zhou et al., 2011 used a computer vision approach called "local binary patterns" that describe pixel neighborhoods in an image (Ojala et al., 1996) to detect slicewise outliers across DWIs, whereas Lauzon et al., 2013 based their detection method to the sum of squared tensor model residuals for each slice. Andersson et al., 2016 proposed the use of a non-parametric Gaussian process (Andersson and Sotiropoulos, 2015) to model and predict diffusion-weighted signals based on the measured DWIs and integrate this with subject motion and distortion corrections (implemented in FSL EDDY). If the measured DWI slice intensities significantly differ from their predicted counterparts, the slice is deemed an outlier and can be replaced by its prediction in further analysis. While this integrated framework provides an elegant solution to the detection of slice outliers before geometrical transformation, the signal prediction sets constraints on the minimal data acquisition and can be time consuming.

In this work, we sought to develop a fast and less complex approach that lifts these restrictions and thus would be compatible with datasets acquired using less gradient directions. To achieve this, we present i) an automated tool coined SOLID (Slicewise OutLIer Detection) to identify slicewise outliers across DWIs before geometrical transformations are applied and ii) a framework to minimize outlier slice impact on estimations which is demonstrated with a tensor model. SOLID is based on a straightforward intensity metric and does not rely on the time-consuming model estimation and predictions of the dMRI signal. Its sensitivity and specificity are extensively evaluated on simulated DWIs with a range of clinical b-values, SNRs, and misalignments of the input images, as well as on 156 data sets from 54 different neonatal subjects. In addition, SOLID is compared to three previously published tools for slicewise outlier detection: DTIPrep (Liu et al., 2010; Oguz et al., 2014), DTI Studio (Jiang et al., 2006; Li et al., 2013), and FSL's EDDY (Andersson et al. 2016, 2017; Andersson and Sotiropoulos, 2016). Instead of removing the entire DWI or replacing slicewise outliers with predictions, we propose here to incorporate the measurement 'uncertainty' derived from SOLID into the estimation (Knutsson and Westin, 1993; Tax et al., 2017; Li et al., 2013). Whereas previous work has minimized the effect of outliers by recuding their weight based on their voxelwise model residuals (Mangin et al., 2001; Meer et al., 1991), here the reduction is based on the combined effect of the brain voxel intensities within the outlier slice without the voxelwise modelling.

2. Methods

The SOLID framework consists of two parts: i) slicewise outlier detection that is based on differences in DWI intensity histograms between slices across DWIs and ii) model estimation informed by the detection results. Fig. 1 gives a schematic overview of the different steps of the SOLID framework. First, we revise the statistical methods of outlier detection based on the Z-score and its robust modified counterpart (Iglewicz and Hoaglin, 1993; Norman and Streiner, 2007) and explain how this was implemented in SOLID. Second, we detail how the information from SOLID can be integrated into a model estimation framework. Third, we present the simulation experiments and real data acquisitions for the evaluation of the detection and model estimation, and finally provide methodological details of the tools that are evaluated for comparison.

2.1. Slicewise OutLIer Detection: SOLID

To examine the likelihood of a slice k of DWI l being an outlier, it is useful to design a slicewise summary feature or 'observation' $y_{k,l}$ that can be compared across slices and/or DWIs. The Z-score can subsequently be used as a statistical approach to identify outliers from the multiple observations y of a random variable Y which are drawn from a normal distribution with a mean μ and a standard deviation σ i.e. $Y \in N(\mu, \sigma^2)$. The steps of SOLID are as follows (Fig. 1A–C, note that this is on the raw non-registered data):

A) The intensity histograms of normal and outlier slices have distinct characteristics (Scelfo et al., 2012).
B) The summary statistic, or the 'observation' y, is a slicewise intensity metric calculated within a brain mask for each slice k and each DWI l per shell, denoted by $y_{k,l}$. In the remainder of the manuscript, we use the variance of the intensities as the metric, but other choices are also possible and are further elaborated upon in the Discussion.
C) Using the Z-score, an observation can be deemed an outlier if the difference between the measurement and the sample mean $\bar{y}$ divided by the sample standard deviation s is large. More specifically, in the case of slicewise outlier detection, the mean $\bar{y}_k$ and the standard deviation s_k can be calculated for each slice k across DWIs per shell, resulting in a slice- and DWI-specific Z-score $z_{k,l}$ (eq. (1)):

$$z_{k,l} = \frac{y_{k,l} - \bar{y}_k}{s_k}. \quad (1)$$

A challenge when using the mean and standard deviation of observations for outlier detection is that in the presence of multiple outliers, a masking effect might occur, and milder outliers could remain undetected. The modified Z-score (Iglewicz and Hoaglin, 1993) was introduced to overcome this drawback by replacing the mean $\bar{y}$ with the median $\tilde{y}$ and the standard deviation s with the median absolute deviation (MAD) (Mangin et al., 2002). The modified Z-score $\tilde{z}_{k,l}$ is calculated with the help of the slicewise median $\tilde{y}_k$ and MAD_k:

$$\tilde{z}_{k,l} = \frac{y_{k,l} - \tilde{y}_k}{MAD_k}, \quad (2)$$

Fig. 1. SOLID framework: the detection of slicewise outliers from DWIs and the informed model estimation. A-B) An intensity metric y (here variance) is computed across all voxels x for each slice k in DWI l, denoted by $y_{k,l}$. C) The modified Z-score $_{k,l}$ is calculated from $y_{k,l}$, the slicewise median $\bar{y}_k$ and the median absolute deviation MAD_k. The color scale already reveals some outlier slices with high values. D) Geometrical misalignment correction step gives image transformation matrices T_l for each volume. Red dashed lines visualize the axial plane before and after transformation relative to the signal decrease slice outlier. E1) Voxelwise interpolated modified Z-scores $_l'(x')$ are obtained using the same transformations T_l for each modified Z-score volume $_l(x)$. E2) A voxelwise SOLID weight $S_l(x')$ is linearly interpolated from the $_l'(x')$ to obtain the certainties of data points. E3) The SOLID-informed model estimation.

$$\mathrm{MAD}_k = 1.4826\ \mathrm{median}\left(\left|y_{k,l} - \bar{y}_k\right|\right). \qquad (3)$$

The slicewise modified Z-score map (Fig. 1C) provides a quick but very useful overview of the data: suspicious slices with a high modified Z-score can be readily displayed for visual inspection. To do actual outlier 'detection and rejection' to reduce the impact on model fitting, one could straightforwardly set an arbitrary threshold on the modified Z-scores and subsequently discard the DWI intensities with the modified Z-score exceeding this threshold during estimation. However, the geometric transformation step to correct for motion and distortion prior to model estimation results in gradual intensity bands (Fig. 1D), which challenges the choice of an appropriate threshold setting because intensities in voxels that are only partially affected by outliers could completely be rejected depending on the threshold setting. Even though the SOLID output allows the user to detect and reject outlier slices in this way, we propose not to adopt this approach but instead to downweight data points based on the degree of partial voluming with the outlier data. This SOLID-based estimation after geometric transformation is described in the next section.

2.2. SOLID-informed estimation after geometric transformation

Incorporating the degree to which a voxel is affected by an outlier measurement to the model estimation is in line with the signal/certainty philosophy, i.e. naturally separating the data into a signal and certainty part (Knutsson and Westin, 1993). Here, we propose to derive the estimate for 'certainty' from interpolated modified Z-scores. We will refer to this certainty as 'SOLID-weight'. The steps of the SOLID-based estimation are as follows (Fig. 1E, note that geometric transformation can be done with any preferred method that can output the transformation for each DWI l (T_l), even in a slice-to-volume manner (Ferrante and Paragios, 2017)):

E1) Transforming the slicewise modified Z-scores $_{k,l}$ into voxelwise SOLID weights after geometric transformation requires first to map them to 3D volumes $_l$ ($l = 1, \ldots, n$ with n being the number of

DWIs) so that all the voxels in the slice k of volume l have the same value, i.e. $\mathbf{Z}_l(i,j,k) = \tilde{Z}(k,l)$. Here, i and j denote the spatial coordinates in the slice-plane. Note that the 3D volumes $\mathbf{Z}_l(i,j,k)$ contain the same information as the modified Z-score map $\tilde{Z}(k,l)$, but have the same dimensions as the image data. To simplify the notation, the 3D volumes can be written as $\mathbf{Z}_l(\mathbf{x})$ with $\mathbf{x}$ being the voxel coordinate. Subsequently, the same volumetric transformations T_l that are used to correct DWIs for geometric distortions are applied to the 3D volumes $\mathbf{Z}_l(\mathbf{x})$, resulting in voxelwise transformed and interpolated modified Z-score maps $\mathbf{Z}'_l(\mathbf{x}') = \mathbf{Z}_l(T_l(\mathbf{x}))$.

E2) Voxelwise SOLID weights $S_l(\mathbf{x}')$ in the transformed image space are derived from the modified Z-scores $\mathbf{Z}'_l(\mathbf{x}')$ by scaling them linearly between 0 (outlier) and 1 (reliable data point) using manually chosen lower and upper thresholds t_L and t_U as shown in eq. (4). In contrast to the voxelwise Geman-McClure weights (Meer et al., 1991; Mangin et al., 2002), SOLID weights are based on the data from all the brain voxels within a slice thus providing additional statistical power for the outlier detection, and do not depend on the residuals of an estimated model.

$$S_l(\mathbf{x}') = \begin{cases} 0 & ; \text{ if } \mathbf{Z}'_l(\mathbf{x}') > t_U \\ \dfrac{\mathbf{Z}'_l(\mathbf{x}') - t_U}{t_L - t_U} & ; \text{ if } t_L \le \mathbf{Z}'_l(\mathbf{x}') \le t_U \\ 1 & ; \text{ if } \mathbf{Z}'_l(\mathbf{x}') < t_L \end{cases} \quad (4)$$

Thresholds $t_U > t_L \ge 0$ can be tuned to adjust for the normal variation of the modified Z-scores within the data. When variance is chosen as the slicewise intensity metric $y_{k,l}$, we recommend a lower threshold of $t_L = 3.5$ based on our experiments and the previous literature (Iglewicz and Hoaglin, 1993). Instead of normalizing to the maximum modified Z-score in $\mathbf{Z}'_l(\mathbf{x}')$ to obtain a SOLID weight between 0 and 1, we propose the use of an upper threshold $t_U \ll \max(\mathbf{Z}'_l(\mathbf{x}'))$ beyond which the SOLID weight is clipped to 1. We found a value two to three times larger than the lower threshold (by default $t_U = 10$) appropriate to downweight outlier data points, while preventing extreme outliers with large modified Z-score from dominating the process. Practically, this means that $\mathbf{Z}'_l(\mathbf{x}') < t_L$ will have a weight of 1, $\mathbf{Z}'_l(\mathbf{x}') > t_U$ will have a weight of 0, and the remaining $\mathbf{Z}'_l(\mathbf{x}')$ will have a value between 0 and 1 during the model estimation; i.e. the edge regions of gradual intensity bands are naturally downweighted based on the degree of partial voluming with the outlier slice.

E3) The SOLID weights are subsequently used to inform a model estimator, where different approaches can be taken depending on the model and the estimator. For the iteratively re-weighted linear least squares (IWLLS) estimator (Veraart et al., 2013), a straightforward incorporation is to multiply the diagonal of the weight matrix by the corresponding SOLID weights $S(\mathbf{x}')$ combined from all shells $S_l(\mathbf{x}')$ during each iteration step m giving weights estimates $\mathbf{W}_m(\mathbf{x}')$ for each voxel $\mathbf{x}'$ based on the design matrix $\mathbf{X}$ and voxelwise tensor element estimates $\hat{\mathbf{b}}(\mathbf{x}')$:

$$\mathbf{W}_m(\mathbf{x}') = \mathrm{diag}\left(S(\mathbf{x}')\right)\exp\left(2\mathbf{X}\hat{\mathbf{b}}(\mathbf{x}')_{m-1}\right). \quad (5)$$

2.3. Simulation experiments

Simulations were based on a part of a human full-brain dataset of the MASSIVE database (Froeling et al., 2017). The data were acquired on a single subject in eight sessions and consisted of 170 non-diffusion weighted images and 250, 500, and 500 images with b-values of 500 s/mm^2, 1000 s/mm^2, and 2000s/mm^2, respectively. A ground truth (GT) dataset was obtained by anisotropically smoothing the data (full width at half maximum of 6 mm) (Leemans et al., 2009) and subsequently estimating diffusion tensor (DT) and kurtosis tensor (KT) using the IWLLS estimator in ExploreDTI (Leemans et al., 2009; Veraart et al., 2011). Finally, the estimated tensors were used to simulate the GT data with non-diffusion-weighted signal and diffusion signal on two shells with b-values of 1000 s/mm^2 and 2000 s/mm^2. The number of diffusion gradient directions per shell is detailed in Table 1.

2.3.1. Outlier detection

To evaluate the sensitivity and specificity of SOLID, receiver operating characteristic (ROC) curves were generated by setting hypothetical thresholds on the modified Z-score values $\tilde{Z}_{k,l}$ and counting the true/false positive/negative slices. A similar approach was used to evaluate positive predictive value through precision – recall curves (PRC). Note that this threshold is ‘hypothetical’ because in the data/certainty approach, a positive finding does not directly lead to the exclusion of the measurement. In all outlier detection simulations, each shell consisted of the same 30 diffusion gradient orientations (Jones et al., 1999). Outlier slices were introduced in DWIs in different setups detailed in Table 1. Positions of the outlier slices were fixed across all affected DWIs to maximize the effect. For each setup, both a complete signal loss (−100%) and a modest signal increase (+50%) artefacts were generated using four SNR levels (8, 16, 32, and infinite) of Rician noise. The random selection of artefactual DWIs and slicewise outliers was repeated to form 1000 unique sets of 30 DWIs for each experiment.

To study whether initial geometrical misalignments of structures resulting from subject motion were problematic for the slicewise comparison in SOLID, 5-degree rotations around the left-right and the anterior-posterior axes were applied to the artefactual DWIs and the experiments described above were repeated. A common brain mask calculated as a union of all DWI brain masks was used in outlier detection simulations.

2.3.2. Comparison with existing algorithms for slicewise outlier detection

SOLID is compared to three previously published tools for slicewise outlier detection: DTIPrep (Liu et al., 2010; Oguz et al., 2014), DTI Studio (Jiang et al., 2006; Li et al., 2013), and FSL's EDDY (Andersson et al., 2016; Andersson and Sotiropoulos, 2016). For the former two tools, we qualitatively compare the results on a simulated dataset whereas for FSL EDDY we did a more complete ROC and PRC profile-based comparison.

DTIPrep (Liu et al., 2010) uses a voxelwise normalized cross-correlation metric between successive slices within each DWI. The algorithm assumes that normalized cross-correlation values calculated at the same slice pair positions across DWIs follow normal distributions. If the normalized cross-correlation score of two successive DWI slices deviates more than 3.5 standard deviations from the mean normalized cross-correlation score at a given slice position, the pair is considered to have an outlier slice. This method by design does not identify which one of the adjacent slices was the outlier. To circumvent this problem, DTIPrep rejects the whole DWI volume if it contains a user-specified number of outlier slices.

DTI Studio (Li et al., 2013) performs a morphological closing operation perpendicular to the acquisition plane and subtracts the result from the original DWI. If there is a slicewise signal decrease outlier, it will get ‘closed’, and the subtraction image shows bright voxels in that area. DTI Studio takes a similar data/certainty approach as proposed here, by estimating the certainty with “corrected inter-slice intensity discontinuity” metric and using it as a factor in model estimation weights.

FSL's EDDY (Andersson and Sotiropoulos, 2016; Andersson et al., 2017) algorithm makes slicewise predictions based on a Gaussian process and compares these predictions with the acquired slices. The detection criterion assumes that voxelwise intensity differences between measured and predicted slices follow a normal distribution. Slices with a difference larger than a predefined number of standard deviations from the mean, are considered outliers. We performed comparison for the ROC curve and the PRC profile by varying the standard deviation threshold setting for

EDDY from 0 to 300 and the modified Z-score threshold for SOLID from 0 to 300. The test data consisted of 30 DWIs with $b = 1000$ s/mm^2 of which eight were randomly replaced with artefactual volumes that had five randomly located outlier slices with either a 100% signal decrease or a 50% signal increase and 5-degree rotation around left-right axis. The selection of artefactual DWIs, outlier slices, and the addition of Rician noise with SNR 16 was repeated 1000 times.

2.3.3. SOLID-informed estimation: the effect of introducing uncertainty

The small degrees of partial voluming with outlier slices or small signal deviations do not significantly affect model estimates. Accompanying the measurement with a certainty (SOLID weight) allows for reducing the effect of different degrees of signal deviation. To evaluate the effect of the degree of signal deviation and the downweighting on diffusion measure estimates, voxels in the corpus callosum (CC) and in the deep gray matter (GM) were selected from the GT to represent typical tissue signals. The CC voxel had fractional anisotropy (FA), mean diffusivity (MD), mean kurtosis (MK), kurtosis anisotropy (KA), and radial kurtosis (RK) (Poot et al., 2010) values of 0.85, 0.87×10^{-3} mm^2/s, 1.5, 0.90, and 3.2, respectively. Corresponding values for GM were 0.15, 0.94×10^{-3} mm^2/s, 0.75, 0.10, and 0.85, respectively. Diffusion signals were simulated on two shells with a 60 gradient direction scheme (Jones et al., 1999) with diffusion weightings of 1000 s/mm^2 and 2000 s/mm^2 along with five non-diffusion weighted signals.

Diffusion signals were simulated with different diffusion weightings detailed in Table 1. Multiple setups were generated by randomly selecting a varying amount of directions (up to 16 out of 60) and introducing signal deviations from a complete signal loss (−100%) to a signal increase (+50%) with 10% steps. This setup encompasses situations where outliers with different signal deviations are fully or partially interpolated into the result. The degree of the signal deviation in each setup was the same for all the directions. For each setup, 1000 Rician noise iterations were introduced with different SNRs, and the DKI equation was fitted using IWLLS (Veraart et al., 2013). The certainty (SOLID weight) was gradually reduced for the outlier signals from 100% to 0% with 10% steps (eq. 5).

Finally, we compare the performance of SOLID-informed IWLLS tensor estimation with normal IWLLS tensor estimation and REKINDLE (Tax et al., 2015) in gradual intensity band regions.

2.4. Real data experiments

The usage of neonatal data was approved by the Ethics Committee of the Helsinki University Hospital and the usage of the adult data was approved by the University of Helsinki Ethical Review Board of Humanities and Social and Behavioural Sciences. All subjects or parents of the newborn subjects had given their written consent.

The sensitivity and specificity of SOLID outlier detection was evaluated with the data of 54 neonatal subjects scanned as a part of a previous study (Stjerna et al., 2015). Full-brain axial DWIs were acquired using Philips Achieva 1.5T scanner (Best, The Netherlands) with an 8-channel phased array head coil. A single-shot echo planar imaging sequence with an acquisition matrix of 128×128, voxel size of $1.75 \times 1.75 \times 2.0$ mm^3 and TR/TE of 6700/58 ms was used to acquire 15 DWIs based on over-plus gradient scheme with b-value of 700 ss1(sca8 0 uviatead)-24usion

Jones, D.K., Horsfield, M.A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn. Reson. Med. 42 (3), 515–525.
Kennis, M., et al., 2016. Choosing the polarity of the phase-encoding direction in diffusion MRI: does it matter for group analysis? Neuroimage: Clinica 11, 539–547.
Knutsson, H., Westin, C., 1993. Normalized and differential convolution: methods for interpolation and filtering of incomplete and uncertain data. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June(16–19), pp. 515–523.
Langen, M., et al., 2012. Fronto-striatal circuitry and inhibitory control in autism: findings from diffusion tensor imaging tractography. Cortex 48 (2), 183–193.
Lauzon, C.B., et al., 2013. Simultaneous analysis and quality assurance for diffusion tensor imaging. PLoS One 8 (4).
Leemans, A., et al., 2009. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine 17 (2), 3537.
Li, Y., et al., 2013. Image corruption detection in diffusion tensor imaging for post-processing and real-time monitoring K. Herholz. PLoS One 8 (10), e49764.
Liu, J., Gasbarra, D., Railavo, J., 2016. Fast estimation of diffusion tensors under Rician noise by the EM algorithm. J. Neurosci. Meth. 257, 147–158.
Liu, Z., et al., 2010. Quality control of diffusion weighted images. In: Proceedings of SPIE–the International Society for Optical Engineering. NIH Public Access, p. 76280J.
Lucas, B.C., et al., 2010. The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software. Neuroinformatics 8 (1), 5–17.
Mangin, J.F., et al., 2002. Distortion correction and robust tensor estimation for MR diffusion imaging. Med. Image Anal. 6, 191–198.
Mangin, J.F., et al., 2001. Eddy-current distortion correction and Robust tensor estimation for MR diffusion imaging. Lect. Notes Comput. Sci. 2208, 186–194.
Marami, B., Scherrer, B., Afacan, O., Erem, B., et al., 2016a. Motion-robust diffusion-weighted brain MRI reconstruction through slice-level registration-based motion tracking. IEEE Trans. Med. Imag. 35 (10), 2258–2269.
Marami, B., Scherrer, B., Afacan, O., Warfield, S.K., et al., 2016b. Motion-robust reconstruction based on simultaneous multi-slice registration for diffusion-weighted MRI of moving subjects. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Cham, pp. 544–552.
Meer, P., et al., 1991. Robust regression methods for computer vision: a review. Int. J. Comput. Vis. 6 (1), 59–70.
Morris, D., et al., 2011. Preterm neonatal diffusion processing using detection and replacement of outliers prior to resampling. Magn. Reson. Med. 66 (1), 92–101.
Nielsen, J.F., Ghugre, N.R., Panigrahy, A., 2004. Affine and polynomial mutual information coregistration for artifact elimination in diffusion tensor imaging of newborns. Magn. Reson. Imag. 22 (9), 1319–1323.
Norman, G.R., Streiner, D.L., 2007. Biostatistics: the Bare Essentials, PMPH-USA.
Odish, O.F.F., et al., 2015. Microstructural brain abnormalities in Huntington's disease: a two-year follow-up. Hum. Brain Mapp. 36 (6), 2061–2074.
Oguz, I., et al., 2014. DTIPrep: quality control of diffusion-weighted images. Front. Neuroinf. 8, 4.
Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recogn. 29 (1), 51–59.
Pannek, K., et al., 2015. Combined slicewise and voxelwise outlier detection in diffusion MRI data of preterm born infants. Ohbm.
Pannek, K., et al., 2012. HOMOR: higher order model outlier rejection for high b-value MR diffusion data. Neuroimage 63 (2), 835–842.
Pierpaoli, C., 2010. Artifacts in diffusion MRI. In Diffusion MRI: Theory, methods, and applications 303–317.
Pierpaoli, C., Walker, L., 2010. TORTOISE: an integrated software package for processing of diffusion MRI data. Processing of Diffusion … 51 (1), 2010.
Poot, D.H.J., et al., 2010. Optimal experimental design for diffusion kurtosis imaging. IEEE Trans. Med. Imag. 29 (3), 819–829.
Rohde, G.K., et al., 2004. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. Magn. Reson. Med. 51 (1), 103–114.
Sairanen, V., et al., 2017. A novel measure of reliability in Diffusion Tensor Imaging after data rejections due to subject motion. Neuroimage 147, 57–65.
Scelfo, D., et al., 2012. Automated detection, evaluation and removal of DWI-related artifacts. In Proc. Intl. Soc. Mag. Reson. Med 3553.
Skare, S., et al., 2000. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. J. Magn. Reson. 147 (2) pp.340–52.
St-Jean, S., Coupé, P., Descoteaux, M., 2016. Non Local Spatial and Angular Matching: enabling higher spatial resolution diffusion MRI datasets through adaptive denoising. Med. Image Anal. 32, 115–130.
Stejskal, E.O., Tanner, J.E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J. Chem. Phys. 42 (1), 288–292.
Stjerna, S., et al., 2015. Visual fixation in human newborns correlates with extensive white matter networks and predicts long-term neurocognitive development. J. Neurosci. 35 (12), 4824–4829.
Tax, C.M.W., et al., 2017. Quantifying the brain's sheet structure with normalized convolution. Med. Image Anal. 39, 162–177.
Tax, C.M.W., et al., 2015. REKINDLE: Robust extraction of kurtosis INDices with linear estimation. Magn. Reson. Med. vol. 73 (2), 794–808.
Tax, C.M.W., Vos, S.B., Leemans, A., 2016. Checking and correcting DTI data. In: Diffusion Tensor Imaging. Springer, New York, NY, pp. 127–150. New York.
Tournier, J.-D., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. Magn. Reson. Med. 65 (6), 1532–1556.
Tournier, J.D., et al., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage 23 (3), 1176–1185.
Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage 35 (4), 1459–1472.
Tuch, D.S., 2004. Q-ball imaging. Magn. Reson. Med. 52 (6), 1358–1372.
Veraart, J., et al., 2013. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. Neuroimage 81, 335–346.
Veraart, J., Van Hecke, W., Sijbers, J., 2011. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. Magn. Reson. Med. 66 (3), 678–686.
Vos, S.B., et al., 2017. The importance of correcting for signal drift in diffusion MRI. Magn. Reson. Med. 77 (1), 285–299.
Wedeen, V.J., et al., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54 (6), 1377–1386.
Zhou, Z., et al., 2011. Automated artifact detection and removal for improved tensor estimation in motion-corrupted DTI data sets using the combination of local binary patterns and 2D partial least squares. Magn. Reson. Imag. 29 (2), 230–242.